Standardizing Clinical Laboratory Data for the Development of Transferable Computer-Based Diagnostic Programs

P. W. Strike, A. Michaeloudis, and A. J. Green

The existence of systematic differences between test results obtained at different laboratories can compromise the development of generally accessible reference databases for interpretive pathology. We review approaches to the elimination of inter-laboratory bias from pathology test results through the use of standard unit transformations. A general transform procedure is described that will permit laboratories serving a common population to make use of reference data, decision rules, and computer-based interpretive programs developed around a larger clinical database than each of these test centers could amass for themselves.

Additional Keyphrases: analytical bias · clinical decision making · quality-control · reference values · standardization diagnosis with the help of a computer

The Transfer Problem

The interpretation of clinical laboratory test results depends upon the availability of appropriate reference data. If several different laboratories service a common population with respect to the biological distribution of the analyte in health and disease, inter-laboratory differences in assay methodology or procedure may prevent the use of a common reference database. Such local effects, or biases, are particularly evident for ill-specified or heterogeneous analytes, and for functional assay systems where the unit of measurement is defined as an effect within a specified analytical system, rather than in terms of molecular number (mass).

Because test results may only be coherently interpreted within a locally established reference framework, the problem arises as to how we may devise generally accessible reference data, decision rules, or computer-based interpretive programs that may be utilized by test centers not directly involved in their development. This problem is evidenced in the recent literature on the transportability of discriminant functions for hypercalcemia. Holland and Jacobs (1) consider the problems involved in utilizing the discriminant equations of Fraser et al. (2) when the units of measurement for alkaline phosphatase in serum differ from those originally used by Fraser et al. Holland and Jacobs used a formal assay-comparison study to derive conversion factors permitting the translation of their results into units equivalent to those of the method used by Fraser et al. Some complex problems are posed by such an approach, not least of which is the assumption that the original alkaline phosphatase assay method, set up specifically to derive the conversion factors, behaved exactly as it did in the hands of Fraser et al. Statistical problems arise with respect to the techniques used to derive the conversion factors, given the existence of random error in both of the assay systems.

The data-transfer problem bedevils any interpretive program, irrespective of the procedure used to formulate the decision rules, be it formal discriminant analysis in its various manifestations, including the equivalent use of Bayes' theorem (3) (under simplifying assumptions) or the less-formal heuristics embodied in rule-based expert systems (4). The case for attacking this problem through the standardization of test results before such interpretive programs are formulated is a strong one.

Here we propose a relatively simple solution to the data-transfer problem, based on a standard unit transformation. Standard units have been proposed in the past as a means of simplifying the clinical interpretation of laboratory data and have been reviewed in this context (5). The best-known standard unit transform, the SD-unit or Z-transform, is given by:

$$Z_i = \frac{X_i - \mu}{\sigma}$$

where $X_i$ is the test value for a test subject and $\mu$ and $\sigma$ are estimates of the reference population parameters $\mu$ and $\sigma$ (mean and standard deviation) in a healthy reference class. The original test values are transformed into standard units whose distribution is centered on a mean of zero with a standard deviation of 1. Under the assumption of a gaussian reference distribution, the derived $Z_i$ values are interpreted as "healthy" within the distribution range of $-2$ to $+2$.

The gaussian assumption has been dropped by some authors through the use of median ($m$) and percentile range estimates as follows:

$$P_{95} = \frac{X_i - m}{(\frac{1}{4} 95\text{-percentile range})}$$

The transformation of equation 2 is the essential element of Lo and Kellen's "normal quotient unit" (6), which differs from equation 2 only in the use of some scaling constants to provide a "universal" reference percentile range of 80 to 120.

A variant on equation 2 has been proposed by Dybkaer (5) as being "the most straightforward and robust" of the linear transforms available, although no support was offered for this view:

$$P_{68} = \frac{X_i - m}{(\frac{1}{4} 68\text{-percentile range})}$$

We note that these transforms do have a secondary effect upon inter-laboratory bias. Systematic disparities between the raw test results obtained for a common set of test specimens, but analyzed at different laboratories or by different analytical methods, appear to be decreased by linear transformation (6–8).

Methods

Modeling the Transfer Problem

For a given test specimen we have estimates $X_{ij}$ of the
level $\xi$ of analyte present as determined by two different assay procedures ($j = 0$ or $1$) where:

$$X_{i0} = \xi + \epsilon_i \tag{4}$$
$$X_{i1} = \eta_i + \epsilon_i$$
$$\eta_i = \alpha + \beta \cdot \xi_i \tag{5}$$

The terms $\epsilon_i$ represent the random-error components in the measurement procedures as expressed in the reported assay values $X_{ij}$; these errors being assumed to be independent and normally distributed with a constant variance $\sigma^2$ throughout the range of the measurements (homoscedastic error). The homoscedastic error assumption is an expedient of simple modeling at this stage and the consequences of relaxing this assumption will be examined subsequently. The parameters $\alpha$ and $1 - \beta$ represent the relative analytical bias (systematic errors) between the values delivered by the two assay systems, with zero expectations in the absence of such biases.

A fundamental assumption of the modeling that follows is that equation 5 describes the behavior of the assay system across the entire range of the test measurements and is quite independent of the clinical status of the test subjects under study. That this model is a reasonable approximation of the biases arising in most clinical assay comparison studies can be deduced from the generally linear form of published assay comparison data. Our use of the word "relative" in describing the bias terms $\alpha$ and $1 - \beta$ is specifically intended to release the model of expression 4 from any constraints with regard to specifying the absolute accuracy of either assay system, i.e., the term $\xi$ is not necessarily to be interpreted as the "true" level of analyte in the test specimen; it is simply an arbitrary benchmark for describing the systematic disparities between the expected test values for the assay systems under study. This maneuver is essential in the context of assays for complex, ill-defined, or molecularly heterogeneous analytes for which there may be no logical basis for specifying accuracy in any absolute physicochemical sense.

Rewriting the model of equation 5 in terms of the observed test values $X_{ij}$ we obtain a structural errors-in-variables model:

$$X_{i1} = \alpha + \beta \cdot X_{i0} + \epsilon_{i1}^* \tag{6}$$

where

$$\epsilon_{i1}^* = \epsilon_{i1} - (\beta \cdot \epsilon_{i0})$$

An obvious approach to eliminating the bias terms from equation 6 is through the medium of a formal assay-comparison study from which we could derive estimates of $\alpha$ and $\beta$ for the purpose of recalibrating the $X_{i1}$ test results to achieve parity with the $X_{i0}$ test results. We have already noted the potential problems associated with this approach.

Solving the Transfer Problem

The strategy we propose takes advantage of the fact that the analytical biases must be expressed in the clinical reference data obtained for the assay systems in the context of a common reference population. For the assay method 0 we estimate the reference distribution parameters $\mu_0$ and $\sigma^2$ (variance) with the following expectations from standard theory:

$$E[X_{i0}] = \mu_0 \quad \therefore \quad \hat{\mu}_0 \tag{7}$$
$$V[X_{i0}] = \sigma^2 = \sigma_\xi^2 + \sigma_\epsilon^2 \quad \therefore \quad \hat{\sigma}_\xi^2 \tag{8}$$

For the assay method 1 we have:

$$E[X_{i1}] = \mu_1 = \alpha + \beta \mu_0 \quad \therefore \quad \hat{\mu}_1 \tag{9}$$
$$V[X_{i1}] = \sigma^2 = \beta^2 \sigma_\xi^2 + \sigma_\epsilon^2 \quad \therefore \quad \hat{\sigma}_\xi^2 \tag{10}$$

Note that equations 7–10 require no formal assumptions about the distribution of the analyte in the population under study or the distribution of the error terms $\epsilon$. We do assume the terms $\xi$ and $\epsilon$ to be statistically independent. The relative bias components $\alpha$ and $1 - \beta$ translate directly into the clinical reference distribution parameter values. In addition, the population variances $\sigma^2$ are directly inflated by the random error variance components of the assay systems.

If the observed test values $X_{i0}$ and $X_{i1}$ are re-scaled by use of the $Z$-transform of equation 1, the relationship between the transformed values is obtained as follows:

Given:

$$Z_{i1} = \frac{X_{i1} - \bar{\mu}_1}{\hat{\sigma}_1} \quad \text{and} \quad Z_{i0} = \frac{X_{i0} - \mu_0}{\hat{\sigma}_0}$$

we have (from equations 6–10):

$$Z_{i1} = \frac{(\alpha + \beta \cdot X_{i0} + \epsilon_{i1}^*) - (\alpha + \beta \cdot \mu_0)}{\hat{\sigma}_1}$$

$$= \frac{\beta (X_{i0} - \mu_0) + \epsilon_{i1}^*}{\hat{\sigma}_1}$$

$$= \left[ \frac{\beta \cdot \hat{\sigma}_0}{\hat{\sigma}_1} \cdot \frac{X_{i0} - \mu_0}{\hat{\sigma}_0} \right] + \epsilon_{i1}^*$$

hence

$$Z_{i1} = \left[ \frac{\beta \cdot \hat{\sigma}_0}{\hat{\sigma}_1} \cdot Z_{i0} \right] + \epsilon_{i1}^* \tag{11}$$

The relative bias term $\alpha$ has been eliminated from the relationship of equation 11, but the proportional bias term $\beta$ remains as a part of the coefficient term $(\beta \cdot \hat{\sigma}_0)/\hat{\sigma}_1$; this coefficient cannot be eliminated from equation 11, leaving us with an irreducible source of bias in the relationship between the standard unit values derived from the $Z$-transform of equation 1. The constant will also manifest itself in the related transforms of equations 2 and 3, by the same line of reasoning as that leading to equation 11.

Because the coefficient $(\beta \cdot \hat{\sigma}_0)/\hat{\sigma}_1$ is essentially a function of the error variance components $\hat{\sigma}_\xi^2$ (as they appear in the structure of $\sigma^2$ from equations 8 and 10), the standardization problem can be resolved by eliminating these error variance terms in the transformation procedure itself, as follows:

$$U_{ij} = \frac{X_{ij} - \mu_j}{\sqrt{\hat{\sigma}_\xi^2 + \hat{\sigma}_\epsilon^2}} \tag{12}$$

In order to model the effects of this transform on the relative bias between the results of two assay methods we introduce a simplified notation:

$$\gamma = \sqrt{\hat{\sigma}^2 - \hat{\sigma}_\xi^2} \quad \delta = \sqrt{\hat{\sigma}^2 - \hat{\sigma}_\epsilon^2}$$

and $U_{ij} = \frac{X_{ij} - \mu_i}{\gamma}$ and $U_{i0} = \frac{X_{i0} - \mu_0}{\delta}$

The relationship between the unit values $U_{ij}$ derived from equation 12 is obtained by using the same arguments as those leading to equation 11:

$$U_{i1} = \left[ \frac{\beta \cdot \delta}{\gamma} \cdot U_{i0} \right] + \epsilon_{i1}^* \tag{13}$$

However, the coefficient term of $U_{i0}$ in equation 13 can now be eliminated by taking account of the expectations defined in equations 7–10. Under the assumption that $\hat{\sigma}_\xi^2$
and $\delta^2_j$ ($j = 0, 1$) are consistent estimates of $\sigma^2_j$ and $\sigma^2_0$ we expect, for large sample sizes:

$$
\delta^2 = \delta^2_0 - \delta^2_1 \xrightarrow{n \to \infty} \sigma^2_0 - \sigma^2_1 = \sigma^2
$$

$$
\gamma^2 = \delta^2_1 - \delta^2_1 \xrightarrow{n \to \infty} \sigma^2_1 - \sigma^2_1 = \sigma^2
$$

so that:

$$
\left[ \frac{\beta - \delta}{\gamma} \right] = \left[ \frac{\delta^2_0 - \delta^2_1}{\gamma^2} \right]^{1/2} = 1
$$

(14)

so that, for large sample sizes, equation 13 tends to:

$$
U_{11} = U_{00} + (\varepsilon_1 \gamma) / \gamma
$$

(15)

The general U-transform of equation 12 has produced an unbiased relationship between the transformed test values. We have made no distributional assumptions for the test variable $\xi$ or for the probabilistic interpretation of the units derived from equation 12. Such interpretation is entirely conditional upon the subsequent treatment of these units in the clinical decision process. The U-transform is simply a linear rescaling of the assay values—their distributional form is unchanged. What is important for our purposes is that any interpretive rules developed around such a transformed reference database can be utilized by any other center accessing that (common) population, through the medium of locally standardized test values.

Example

To illustrate the U-transform, let us consider a comparison of two creatine kinase (CK) assay methods, using as subjects 48 healthy non-pregnant women (ages 20 to 43). Figure 1a shows the mean values for CK in serum, based on triplicate assays by each assay system.

Table 1 illustrates the summary statistics required for the U-transform procedure and the steps in transforming the CK values to U values. The error variance estimates were supplied by the authors of the CK study (8), these being adjusted for the replication entailed in the comparison. The

---

*By "consistent" we imply the following:

$$
\delta^2_j \xrightarrow{n \to \infty} \sigma^2_j
$$

$$
\delta^2_0 \xrightarrow{n \to \infty} \sigma^2_0
$$

i.e., as the sample size $n$ tends to infinity we expect the sample estimates $\delta^2_j$ and $\delta^2_0$ to tend to the parameter values $\sigma^2_j$ and $\sigma^2_0$, respectively.

---

**Table 1. Summary Statistics and Illustration of U-Transform Procedure for the CK Assay Data**

<table>
<thead>
<tr>
<th>Test data</th>
<th>Original data</th>
<th>Standardized data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$X_0$</td>
<td>$X_1$</td>
</tr>
<tr>
<td>$i = 1$</td>
<td>2.5</td>
<td>147.6</td>
</tr>
<tr>
<td>$i = 2$</td>
<td>1.4</td>
<td>93.4</td>
</tr>
<tr>
<td>$i = 3$</td>
<td>1.8</td>
<td>113.3</td>
</tr>
<tr>
<td>$i = 4$</td>
<td>1.6</td>
<td>100.4</td>
</tr>
<tr>
<td>$i = 48$</td>
<td>1.9</td>
<td>97.7</td>
</tr>
</tbody>
</table>

$X_1 = 4.807 + 56.489 X_0$

$U_1 = 0.00 + 1.008 U_0$

---

The test data and the transformed data are shown in Figure 1, $a$ and $b$. The relationship between the results of the two CK assay methods before and after transformation has been determined with use of estimators appropriate to the errors-in-variables model of equation 6 (see Appendix A, equations A4–A5).

Raw CK data: $X_0 = 4.81 + 58.49 X_1$

U-transform data: $U_B = 0.00 + 1.008 U_A$

The U-transform has effectively eliminated any systematic disparities between the two sets of CK test results. A decision rule based on the transformed data for one assay will be directly applicable to the interpretation of the transformed results of the other; e.g., the 90th percentile of the transformed reference data for method A is 1.44, while the same value represents the 87.5th percentile for method B, the small difference being well within the bounds of sampling error.

The U-transform will facilitate several objectives:

a. the formulation of transferable rules for the interpretation of test results from different laboratories serving a common population;

b. the pooling of reference data from different laboratories serving a common population; and

c. the standardization of test results from different laboratories serving a common population.

Figure 2 summarizes the general procedure for developing and interpreting rule under the U-transform procedure.

The U-Transform in Detail

A number of points require clarification regarding the general unit transform of equation 12.

i. Because the total random error of an assay method can be partitioned into within-batch and between-batch sources, the exact nature of the error variance term $\sigma^2_0$ used in the U-transform needs to be understood. We define $\sigma^2_0$ as the error variance component embodied in the original estimate of the healthy reference distribution variance $\sigma^2$, implicit in equations 8 and 10. If the test specimens in the original reference distribution study were assayed entirely in one batch (one recalibration of the assay system), the
appropriate error variance term would be the within-batch error variance component. It is more likely, and desirable, that the reference specimens were assayed in more than one batch, in which case the appropriate error variance term will incorporate both within-batch and between-batch errors. Estimation of the appropriate value for the error variance term in this case is quite straightforward if the procedure described by Rodbard (9) is used.

**ii.** The variance terms in equation 12 are sample estimates, and thus subject to random sampling error; thus it is possible that \( \delta_{	ext{est}}^2 \) exceeds \( \delta_{	ext{true}}^2 \), leading to problems with the transformation procedure of equation 12. We minimize this possibility by assuming that the random error variance is less than \( \frac{1}{4} \) of the true variance of the analyte in the healthy reference population, in line with the recommendations of the Aspen Conference on Analytical Goals in Clinical Chemistry (10). We further assume that the variance estimates are based on studies of many samples. As a quite arbitrary rule of thumb we take "many" to mean \( \geq 30 \) for the error variance \( \delta_{	ext{true}}^2 \) and \( \geq 100 \) for the observed healthy reference distribution variance \( \delta_{	ext{est}}^2 \).

**iii.** The residual error term \( e_i \gamma /g \) in equation 15 appears as a ratio of two random variables, which raises some potentially difficult problems if we attempt to specify its distribution—i.e., we have no grounds for assuming normality even in the case of a ratio of two normal variates (which in fact leads to a Cauchy distribution). Notice that the term \( \gamma \) is made up of two variance component estimates derived from a sample of test results for healthy reference subjects. It is in the nature of a reference distribution that it is estimated once only for a given analyte at a given test center—i.e., we do not accompany every biochemical assay with a re-estimate of the reference distribution parameters. The term \( \gamma \) therefore has the status of a constant in equation 16, from which it follows that the distribution of \( e_i \gamma /g \) is determined completely by the distribution of \( e_i \). Under the assumption of independent normally distributed errors \( e_i \), we have a normal distribution for \( e_i \gamma /g \) in equation 6 and hence for the residual error in equation 15:

\[
(e_i \gamma /g) \sim N\left(0, \frac{(\beta^2 \cdot \sigma_{e_i}^2)}{(\sigma_{e_i}^2 - \sigma_{e_{est}}^2)}\right)
\]  

**iv.** It is almost certain that the error variance of an assay system will vary in time, obliging us to draw a distinction between the error variance \( \sigma_{e_i}^2 \) determined from the original reference data, which is actually used in the U-transform process, and the error variance associated with later test measurements that are subjected to U-transformation. Denoting this later error variance term by \( \sigma_{e_{est}}^2 \), we find that the only effect of such a change is in the residual variance in equation 16:

\[
V(e_i \gamma /g) = \frac{(\sigma_{e_i}^2 + \beta^2 \cdot \sigma_{e_{est}}^2)}{(\sigma_{e_i}^2 - \sigma_{e_{est}}^2)}
\]

The changing variance of the assay system will have no effect upon the elimination of the analytical bias terms and \( 1 - \beta \) from equation 15.

**Heteroscedasticity**

The U-transform modeling can be directly generalized to accommodate heteroscedastic error variances, that is, a functional dependence of \( \sigma_{e_i}^2 \) on the level of the test variable \( \xi \), which we represent in general terms as follows:

\[
\sigma_{e_i}^2 = K \cdot (\xi)
\]

We make no assumptions about the exact form of the function \( K \cdot (\xi) \), which may in practice be quite complex. From standard theory we find the expectation of the error variance across the entire reference data set:

\[
E[\sigma_{e_i}^2] = K^2 \cdot (\sigma_0^2 + \mu^2)
\]

Given replicated measurements on every test specimen in the healthy reference group, we can obtain an estimate of this expected variance by pooling the individual variance estimates obtained for each of the \( n \) reference specimens as follows:

\[
\frac{\sum \sigma_{e_i}^2}{n} = (\sigma_0^2 + \mu^2)/(1/n)
\]

Given consistent estimates of the individual variance terms \( \sigma_{e_i}^2 \), the pooled error variance estimate of equation 20 tends in probability to \( K^2 \cdot (\sigma_0^2 + \mu^2) \). Substituting the pooled error variance estimate in the U-transform of equation 12, we derive exactly the same unbiased relationship between the \( U \) values as expressed in equation 15. An example follows in the next section, in which we illustrate the application of the U-transform under heteroscedastic error conditions.

**Real Decisions—Real Problems**

Interpretation of laboratory data is typically a multivariate process. The precise manner in which this information is utilized is a subject of increasing interest from the point of view of computer support. The U-transform will facilitate the transfer of such computer-support irrespective of the exact procedures used, because the information content of the reference database is invariant under simple linear transformation.

Bayes' decision rule and related discriminant analysis
techniques were developed without assumptions as to distribution but are known to be optimal (in the sense of total probability of misclassification) under multivariate normal assumptions (11). This leads us to consider the problems that can arise in combining the U-transform with secondary normalization transforms.

We note in particular the use of Z-unit transform of equation 1 by Boyd and Lacher (12) as a preliminary step in the application of their multi-stage gaussian transformation procedure. The substitution of the U-transform of equation 12 for equation 1 in this procedure means that there is no reason why this general normalization technique cannot be integrated with the development of a generally accessible interpretive database.

To illustrate the problems that can arise in this context, we have used a simulation program (Appendix A) to generate test values from a relatively simple model of urinary 11b-hydroxycorticosteroid (11-OHCS) for 100 obese and (or) hirsute women, the non-Cushing's reference subjects, and for 100 patients with Cushing's syndrome, by each of two 11-OHCS assay methods. The statistical data for the log-normal population distributions were published by Levell (13), these being derived from a study by Mattingley and Tyler (14), who proposed the 11-OHCS variable as the basis of a decision-rule for undertaking further investigation for Cushing's syndrome in subjects whose signs and (or) symptoms were "not sufficiently florid to merit immediate admission for full investigation" (13).

We elected to use a computer simulation to generate the assay-comparison data for this example, in order to observe the U-transform under relatively severe error conditions (σ0 = 0.5σf) and to introduce bias components of known magnitude (1 - β = 0.22 and α = 61.5 nmol/24 h) into the test data. To simulate a simple heteroscedastic error model, we defined the error variance components as proportional functions of the test level ƒ.

The basic test data are presented in Figure 3a, along with the reference mean and variance estimates required for the U-transformation procedure. The pooled estimates of error variance shown in Figure 3c were derived from replicate assays by each assay method of samples from all the reference subjects (Appendix A).

The test results have been transformed into U-values in Figure 3b, effectively eliminating the relative bias components in the relationship between the standardized assay values.

Raw 11 OHCS data: X1 = 61.5 + 0.78 X0
U-transform data: U1 = -0.04 + 0.99 U0

Our treatment of the 11-OHCS test values has produced a high degree of correspondence between the results obtained by each assay system, sufficient to permit the effective interpretation of results from one assay system with use of decision rules based on the results from the other system. In order to determine a formal cutoff value for identifying subjects for further investigation it would be convenient to utilize non-normal distribution theory, but this is impossible as the data stand because the standardized test values retain the log-normal characteristics of the original test data. To satisfy normal-distribution assumptions a secondary transformation would be required.

A simple logarithmic transformation is impossible in view of the negative scaling of the U-values, so we have used a log(U + constant) transform, the constant being determined iteratively to achieve a zero coefficient of skewness. It might appear that the "negative value" problem could be avoided by normalizing the raw test data before applying a U-transformation. A strong objection to this approach is the presence of the marked constant bias (α) in the raw test data, which will be translated into a nonlinear bias component by a log(x) transform. This would leave us with a clear violation of the model assumptions underpinning the U-transform and would severely complicate any subsequent statistical analysis of the data. This problem will arise with any normalizing transformation, given that they all entail a nonlinear treatment of the data. By using the U-transform as a mandatory first step in the analysis we can at least be sure of minimizing any constant bias and the problems that it will generate in subsequent normalizing transformations.

The normalized U-values for each assay method are plotted in Figure 4a and are represented as frequency distributions in figures 4b and c, superimposed on which are the appropriate normal distribution density curves. Using standard normal theory (15) the intersect point of the Cushing's and non-Cushing's distribution curves have been calculated as 0.663 (for assay 0) and 0.663 (for assay 1), the small difference apparent in these values being compatible with sampling error.
If we formulate a simple decision rule using the assay 0 database, i.e., investigate all subjects with urinary 11-OHCS values exceeding 0.663 log_{10} (U_j + 2.8) units, and apply this rule to the transformed test results actually used in this study, we find that 5% of the Cushing's subjects are misclassified by assay method 0 and 5% by assay method 1.

We might contrast this very effective transfer of the decision rule with the consequences of working entirely in terms of the untransformed test values (nmol/L). The intersect point for the Cushing's/non-Cushing's distribution curves when assay method 0 is used is 986 nmol/L. Using this value as a decision rule to classify the original test results from each assay system, we find that 5% of the Cushing's subjects are misclassified by assay method 0 whereas 14% are misclassified by assay method 1, revealing the restricted utility of interpretive rules formulated in the presence of unresolved relative biases.

The Sceptical Chemist

Care is required in applying normalizing transformations in clinical decision/classification problems. Our 11-OHCS example is relatively straightforward, in that the reference and disease groups are both log-normal. There are no grounds for assuming that nature is so disposed; indeed, it is likely that the distribution of an analyte will differ between healthy and disease subjects.

No single mathematical transformation can impose normality on both the healthy and disease group distributions in these circumstances, and attempts to do so may well diminish the discriminatory power of the test variable (12). Before embarking upon the oceans of normalization, check your data for leaks!

The Transfer Problem in Practice: A Cautionary Tale

The United Kingdom Collaborative Study Group on Alpha-fetoprotein (AFP) in Relation to Neural-tube Defects (NTD) (16) provides a national reference database for the interpretation of serum and amniotic-fluid AFP values in healthy and NTD-affected pregnancies. Test data from 19 collaborating centers were pooled after the use of a simple median transformation:

\[ \text{MoM}_j = \frac{X_{ji}}{\bar{m}_j} \]  \hspace{1cm} (21)

where the median \( \bar{m}_j \) was established for each assay center using the local assay system. This "multiple of median" (MoM) transform was reported as being effective in decreasing the otherwise marked interlaboratory variation in AFP test results from the collaborating test centers. The transformed reference database is illustrated in Figure 5. This
standardization procedure permitted the development of a generally accessible risk-assessment program for NTD risk in pregnancy (17) with use of a multivariate database (gestational dates, prior history of NTD-affected pregnancies, standardized serum and amniotic-fluid AFP results, amniotic-fluid acetylcholinesterase screen, and bloodstaining status of the amniotic fluid sample).

The MoM transform has limitations that are not evident in the general U-transform of equation 12.

In the context of the model of equations 4 through 10, the expected relationship between the MoM values from two centers is obtained on using similar arguments as those leading to equation 11:

\[
\text{MoM}_{11} = \frac{\alpha}{\text{m}_1} + \left( \beta \cdot \frac{\text{m}_0}{\text{m}_1} \cdot \text{MoM}_0 \right) + \frac{\text{e}_1}{\text{m}_1} \tag{22}
\]

No further simplification is possible unless we introduce the highly restrictive assumption that the constant bias term \(\alpha = 0\) (given \(\text{m}_1 = \alpha + \beta \cdot \text{m}_0\), in which case equation 22 reduces to:

\[
\text{MoM}_{1} = \text{MoM}_0 + \left( \frac{\text{e}_1}{\text{m}_1} \right) \tag{23}
\]

If \(\alpha \neq 0\), the MoM transform is ineffective as a basis for standardizing test results from different centers. To underline this limitation we have applied the MoM transform to the CK and \(\beta\)-OHCS examples used in this report:

CK assay data: MoM_B = 0.052 + 0.877 MoM_A
11-OHCS assay data: MoM_1 = 0.130 + 0.838 MoM_0

These results conform exactly to the model predictions of equation 22. Because the MoM-transform has been used in the construction of a major national reference database, it is worth noting the severe limitations on its general use for standardizing test data, implicit in equation 22.

Discussion

The general unit-transform procedure is effective in producing comparable unit values under a wide range of sampling conditions and under very severe error bounds for the test data. To be of value in the construction of transferable interpretive programs/decision-making aids, the procedure demands access to sound estimates of local reference population parameters for the analyte(s) concerned. This requirement is a necessary precondition of setting up such programs in the first place. By taking advantage of the general unit-transform procedure before embarking upon the formulation of interpretive rules/decision criteria, locally established diagnostic aids can be rendered generally accessible to any other test center that is servicing that same population. The MoM transform procedure was a specific application of this principle in the context of a multivariate risk-assessment program for neural-tube defects in pregnancy (17). The U-transform generalizes the application of this principle. Being a linear transformation, the proposed procedure involves no degradation of the information content of the reference database.

We thank Prof. M. J. R. Healy for his helpful criticism of a preliminary draft of this paper, and Dr. N.R. Dennis for his permission to use the CK assay data presented in Figure 1. We also acknowledge the helpful comments of a reviewer on the heterocedastic generalization of our model equations, and the Director General of Royal Air Force Medical Services for permission to publish this paper.

Appendix A

Simulation Studies

A computer simulation of the model equations 4–10 was implemented on a Wang 2200-MVP computer system. The program sequence was as follows:

1. Test values \(\delta_i\) were generated from a log-gaussian distribution function for 100 non-Cushing's reference subjects, using published parameter estimates for the 11-OHCS variable (13).

The basic random number generator was called as a BASIC II function and was extensively tested by the authors, using standard techniques, before acceptance for use. As an additional safeguard, a random permutation subroutine was introduced to shuffle the random deviates in blocks of 17, following the recommendation of Hill in the discussion section of Atkinson and Perdue (18).

2. A set of corresponding \(n\) values were formed from the \(\delta_i\) values with use of equation 5 with defined bias parameter values.

3. Gaussian error components were generated for each of the basic test values to form the observed test values \(X_0\) and \(X_1\), following equation 4. The normal deviate subroutine of Marsaglia and Bray (19) was used for this purpose, following the recommendations of Atkinson and Perdue (18). The error variances \(\sigma_{e_1}^2\) and \(\sigma_{e_2}^2\) were specified for each simulation run as first-order functions of \(\delta_i\) to simulate a simple heterocedastic error model.

4. For each complete reference data set, the parameters \(\mu_2\) and \(\sigma_2\) were estimated as:

\[
\mu_2 = \frac{\sum X_{ij}}{n} \tag{A1}
\]

\[
\sigma_2 = \frac{(\sum (X_{ij} - \mu_2)^2)(n - 1)}{n} \tag{A2}
\]

5. Steps 1–3 were then repeated to generate a completely new study sample of 200 test values (100 non-Cushing's subjects and 100 Cushing's subjects) for the purpose of demonstrating the following standard transforms:

\[
\text{a. } \text{MoM}_0 = \frac{X_{ij}}{m_j} \tag{21}
\]

\[
\text{b. } U_{ij} = \frac{X_{ij} - \mu_j}{\sqrt{\sigma_{e_1}^2 - \sigma_{e_2}^2}} \tag{12}
\]

The bias parameters \(\alpha\) and \(\beta\) were estimated for the study sample values, before and after transformation, with use of standard estimates for the structural errors-in-variables model (20) of equation 6, as follows:

\[
\hat{\beta} = \frac{\sigma_2^2 - \lambda \cdot \sigma_{e_2}^2 + 4 \cdot \lambda \cdot \sigma_{e_1}^2}{2 \cdot \sigma_{e_1}^2} \tag{A4}
\]

\[
\hat{\alpha} = \mu_2 - \hat{\beta} \cdot \mu_0
\]

\[
\sigma_{e_1}^2 = \frac{(\sum (X_{ij} - \mu_0)^2)(n - 1)}{n} \tag{A5}
\]

\[
\sigma_{e_2}^2 = \frac{(\sum (X_{ij} - \mu_0)^2)(n - 1)}{n}
\]

\[
\sigma_{e_1}^2 = \frac{\sum (X_{ij} - \mu_2)^2(n - 1)}{n - 1}
\]

\[
\sigma_{e_2}^2 = \frac{\sum (X_{ij} - \mu_0)^2(n - 1)}{n - 1}
\]

\[
\lambda = \frac{\sigma_{e_2}^2}{\sigma_{e_1}^2}
\]

Note that the mean values \(\hat{\mu}_0\) and \(\hat{\mu}_1\) in equations A4 and A5 are based on the study sample values generated in step 5. The error variance estimates used in equation A5 were derived directly from the study sample values by using
equation A3. Two test values $X_y$ were in fact generated for each of the $\xi$ and $\Omega$ values (for both Cushing’s and non-Cushing’s subjects) for this specific purpose.

References